RAW SEQUENCE LISTING
ERROR REPORT
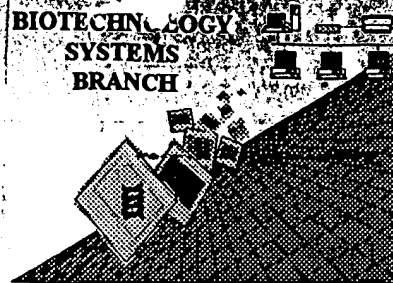
The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/697,206_

Source: _C. PE_

Date Processed by STIC: _11/8/2000_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin30help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

Raw Sequence Listing Error Summary

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 _____ Wrapped Nucleics

The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

2 _____ Wrapped Aminos

The amino acid-number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

3 _____ Incorrect Line Length

The rules require that a line not exceed 72 characters in length. This includes spaces.

4 _____ Misaligned Amino Acid Numbering

The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs
between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

5 _____ Non-ASCII

This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 _____ Variable Length

Sequence(s) _____ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and
indicate in the (ix) feature section that some may be missing.

7 _____ PatentIn ver. 2.0 "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid
sequence(s) _____. Normally, PatentIn would automatically generate this section from the
previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section
to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223>
sections for Artificial or Unknown sequences.

8 _____ Skipped Sequences (OLD RULES)

Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 _____ Skipped Sequences (NEW RULES)

Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

10 _____ Use of n's or Xaa's (NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11 _____ Use of <213>Organism (NEW RULES)

Sequence(s) _____ are missing this mandatory field or its response.

12 _____ Use of <220>Feature (NEW RULES)

Sequence(s) _____ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of new Rules)

13 _____ PatentIn ver. 2.0 "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted
file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).
Instead, please use "File Manager" or any other means to copy file to floppy disk.

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/697,206

DATE: 11/08/2000
TIME: 10:45:55

Input Set : A:\Seqlist.txt
Output Set: N:\CRF3\11082000\I697206.raw

```
 4  <110> APPLICANT: Daniel E.H. Afar
 5        Arthur B. Raitano
 6        Rene S. Hubert
 7        Steve Chappell Mitchell
 8        Aya Jakobovits
10  <120> TITLE OF INVENTION: NOVEL GENE UPREGULATED IN CANCERS OF THE
11        PROSTATE
13  <130> FILE REFERENCE: 129.21-US-U1
15  <140> CURRENT APPLICATION NUMBER: US/09/697,206
15  <141> CURRENT FILING DATE: 2000-10-26
15  <150> PRIOR APPLICATION NUMBER: 60/162,364
16  <151> PRIOR FILING DATE: 1999-10-28
18  <160> NUMBER OF SEQ ID NOS: 26
20  <170> SOFTWARE: FastSEQ for Windows Version 4.0
22  <210> SEQ ID NO: 1
23  <211> LENGTH: 3585
24  <212> TYPE: DNA
25  <213> ORGANISM: Homo Sapiens
27  <400> SEQUENCE: 1
28  cttttggga tcactgctgg ggccaccggg gccaagctag gctcggatga gaaggagttg    60
29  atcctgctgt tctggaaagt cgtggatctg gccaacaaga aggtgggaca gttgcacgaa   120
30  gtgctagtta gaccggatca gttggaactg acggaggact gcaaagaaga aactaaaata   180
31  gacgtcgaaa gcctgtcctc ggcgtcgcag ctggaccaag ccctccgaca gtttaaccag   240
32  tcagtgagca atgaactgaa tattggagta gggacttcct tctgtctctg tactgatggg   300
33  cagcttcatg tcaggcaaat cctgcatcct gaggcttcca agaagaatgt actattacct   360
34  gaatgcttct attccttttt tgatcttcga aaagaattca agaaatgttg ccctggttca   420
35  cctgatattg acaaactgga cgttgccaca atgacagagt atttaaattt tgagaagagt   480
36  agttcagtct ctcgatatgg agcctctcaa gttgaagata tggggaatat aattttagca   540
37  atgatttcag agccttataa tcacaggttt tcagatccag agagagtgaa ttacaagttt   600
38  gaaagtggaa cttgcagcaa gatggaactt attgatgata acaccgtagt cagggcacga   660
39  ggtttaccat ggcagtcttc agatcaagat attgcaagat tcttcaaagg actcaatatt   720
40  gccaagggag gtgcagcact ttgtctgaat gctcaggttc gaaggaacgg agaagctctg   780
41  gttaggtttg taagtgagga gcaccgagac ctagcactac agaggcacaa acatcacatg   840
42  gggacccggt atattgaggt ttacaaagca acaggtgaag atttccttaa aattgctggt   900
43  ggtacttcca atgaggtagc ccagtttctc tccaaggaaa atcaagtcat tgttcgcatg   960
44  cggggggctcc ctttcacggc cacagctgaa gaagtggtgg ccttcttttg acagcattgc  1020
45  cctattactg ggggaaagga aggcatcctc tttgtcacct acccagatgg taggccaaca  1080
46  ggggacgctt ttgtcctctt tgcctgtgag gaatatgcac agaatgcgtt gaggaagcat  1140
47  aaagacttgt tgggtaaaag atacattgaa ctcttcagga gcacagcagc tgaagttcag  1200
48  caggtgctga atcgattctc ctcggccccct ctcattccac ttccaacccc tcccattatt  1260
49  ccagtactac ctcagcaatt tgtgcccccct acaaatgtta gagactgtat acgccttcga  1320
50  ggtcttccct atgcagccac aattgaggac atcctggatt tcctggggga gttcgccaca  1380
51  gatattcgta ctcatggggt tcacatggtt ttgaatcacc agggccgccc atcaggagat  1440
52  gcctttatcc agatgaagtc tgcggacaga gcattatgg ctgcacagaa gtgtcataaa   1500
53  aaaaacatga aggacagata tgttgaagtc tttcagtgtt cagctgagga gatgaacttt  1560
54  gtgttaatgg ggggcacttt aaatcgaaat ggcttatccc caccgccatg cctgtctcct  1620
55  ccctcctaca catttccagc tcctgctgca gttattccta cagaagctgc catttaccag  1680
```

```
56 ccctctgtga ttttgaatcc acgagcactg cagccctcca cagcgtacta cccagcaggc   1740
57 actcagctct tcatgaacta cacagcgtac tatcccagcc ccccaggttc gcctaatagt   1800
58 cttggctact tccctacagc tgctaatctt agcggtgtcc ctccacagcc tggcacggtg   1860
59 gtcagaatgc agggcctggc ctacaatact ggagttaagg aaattcttaa cttcttccaa   1920
60 ggttaccagt atgcaaccga ggatggactt atacacacaa atgaccaggc caggactcta   1980
61 cccaaagaat gggtttgtat ttaagggccc cagcagttag aacatcctca gaaaagaagt   2040
62 gtttgaaaga tgtatggtga tcttgaaacc tccagacaca agaaaacttc tagcaaattc   2100
63 aggggaagtt tgtctacact caggctgcag tattttcagc aaacttgatt ggacaaacgg   2160
64 gcctgtgcct tatcttttgg tggagtgaaa aaatttgagc tagtgaagcc aaatcgtaac   2220
65 ttacagcaag cagcatgcag catacctggc tctttgctga ttgcaaatag gcatttaaaa   2280
66 tgtgaatttg gaatcagatg tctccattac ttccagttaa agtggcatca taggtgtttc   2340
67 ctaagtttta agtcttggat aaaaactcca ccagtgtcta ccatctccac catgaactct   2400
68 gttaaggaag cttcattttt gtatattccc gctcttttct cttcatttcc ctgtcttctg   2460
69 cataatcatg ccttcttgct aagtaattca agcataagat cttggaataa taaaatcaca   2520
70 atcttaggag aaagaataaa attgttattt tcccagtctc ttggccatga tgatatctta   2580
71 tgattaaaaa caaattaaat tttaaaacac ctgaagataa attagaagaa attgtgcacc   2640
72 ctccacaaaa catacaaagt ttaaaagttt ggatcttttt ctcagcaggt atcagttgta   2700
73 aataatgaat tagggggccaa aatgcaaaac gaaaaatgaa gcagctacat gtagttagta   2760
74 atttctagtt tgaactgtaa ttgaatattg tggcttcata tgtattattt tatattgtac   2820
75 ttttttcatt attgatggtt tggactttaa taagagaaat tccatagttt ttaatatccc   2880
76 agaagtgaga caatttgaac agtgtattct agaaaacaat acactaactg aacagaagtg   2940
77 aatgcttata tatattatga tagccttaaa ccttttttcct ctaatgcctt aactgtcaaa   3000
78 taattataac cttttaaagc ataggactat agtcagcatg ctagactgag aggtaaacac   3060
79 tgatgcaatt agaacaggta ctgatgctgt cagtgtttaa cactatgttt agctgtgttt   3120
80 atgctataaa agtgcaatat tagacactag ctagtactgc tgcctcatgt aactccaaag   3180
81 aaaacaggat ttcattaagt gcattgaatg tggatatttc tctaagttac tcatattgtc   3240
82 ctttgcttga atgcaatgcc gtgcagattc atgaggctgc tatttttatt ttctgtgcat   3300
83 tactttaaca ccttaaaggg agaagcaaac atttccttct tcagctgact ggcaatggcc   3360
84 cttttaactgc aataggaaga aaaaaaaaaa ggtttgtgtg aaaattggtg ataactggca   3420
85 cttaagatcg aaaagaaatt tctgtatact tgatgcctta agatgcccaa agctgcccaa   3480
86 agctctgaaa gactttaaga taggcagtaa tgcttactac aatactactg agtttttgta   3540
87 gagttaacat ttgataataa aacttgcctg tttaatctca aaaaa                    3585
89 <210> SEQ ID NO: 2
90 <211> LENGTH: 517
91 <212> TYPE: PRT
92 <213> ORGANISM: Homo Sapiens
94 <400> SEQUENCE: 2
95 Met Thr Glu Tyr Leu Asn Phe Glu Lys Ser Ser Ser Val Ser Arg Tyr
96 1               5                   10                  15
97 Gly Ala Ser Gln Val Glu Asp Met Gly Asn Ile Ile Leu Ala Met Ile
98             20                  25                  30
99 Ser Glu Pro Tyr Asn His Arg Phe Ser Asp Pro Glu Arg Val Asn Tyr
100         35                  40                  45
101 Lys Phe Glu Ser Gly Thr Cys Ser Lys Met Glu Leu Ile Asp Asp Asn
102     50                  55                  60
103 Thr Val Val Arg Ala Arg Gly Leu Pro Trp Gln Ser Ser Asp Gln Asp
104 65                  70                  75                  80
105 Ile Ala Arg Phe Phe Lys Gly Leu Asn Ile Ala Lys Gly Gly Ala Ala
106             85                  90                  95
```

RAW SEQUENCE LISTING   DATE: 11/08/2000
PATENT APPLICATION: US/09/697,206   TIME: 10:45:55

Input Set : A:\Seqlist.txt
Output Set: N:\CRF3\11082000\I697206.raw

```
107 Leu Cys Leu Asn Ala Gln Gly Arg Arg Asn Gly Glu Ala Leu Val Arg
108             100             105             110
109 Phe Val Ser Glu Glu His Arg Asp Leu Ala Leu Gln Arg His Lys His
110         115             120             125
111 His Met Gly Thr Arg Tyr Ile Glu Val Tyr Lys Ala Thr Gly Glu Asp
112     130             135             140
113 Phe Leu Lys Ile Ala Gly Gly Thr Ser Asn Glu Val Ala Gln Phe Leu
114 145             150             155             160
115 Ser Lys Glu Asn Gln Val Ile Val Arg Met Arg Gly Leu Pro Phe Thr
116             165             170             175
117 Ala Thr Ala Glu Glu Val Val Ala Phe Gly Gln His Cys Pro Ile
118         180             185             190
119 Thr Gly Gly Lys Glu Gly Ile Leu Phe Val Thr Tyr Pro Asp Gly Arg
120         195             200             205
121 Pro Thr Gly Asp Ala Phe Val Leu Phe Ala Cys Glu Glu Tyr Ala Gln
122     210             215             220
123 Asn Ala Leu Arg Lys His Lys Asp Leu Leu Gly Lys Arg Tyr Ile Glu
124 225             230             235             240
125 Leu Phe Arg Ser Thr Ala Ala Glu Val Gln Gln Val Leu Asn Arg Phe
126             245             250             255
127 Ser Ser Ala Pro Leu Ile Pro Leu Pro Thr Pro Pro Ile Ile Pro Val
128             260             265             270
129 Leu Pro Gln Gln Phe Val Pro Pro Thr Asn Val Arg Asp Cys Ile Arg
130         275             280             285
131 Leu Arg Gly Leu Pro Tyr Ala Ala Thr Ile Glu Asp Ile Leu Asp Phe
132     290             295             300
133 Leu Gly Glu Phe Ala Thr Asp Ile Arg Thr His Gly Val His Met Val
134 305             310             315             320
135 Leu Asn His Gln Gly Arg Pro Ser Gly Asp Ala Phe Ile Gln Met Lys
136             325             330             335
137 Ser Ala Asp Arg Ala Phe Met Ala Ala Gln Lys Cys His Lys Lys Asn
138             340             345             350
139 Met Lys Asp Arg Tyr Val Glu Val Phe Gln Cys Ser Ala Glu Glu Met
140         355             360             365
141 Asn Phe Val Leu Met Gly Gly Thr Leu Asn Arg Asn Gly Leu Ser Pro
142     370             375             380
143 Pro Pro Cys Leu Ser Pro Pro Ser Tyr Thr Phe Pro Ala Pro Ala Ala
144 385             390             395             400
145 Val Ile Pro Thr Glu Ala Ala Ile Tyr Gln Pro Ser Val Ile Leu Asn
146             405             410             415
147 Pro Arg Ala Leu Gln Pro Ser Thr Ala Tyr Tyr Pro Ala Gly Thr Gln
148             420             425             430
149 Leu Phe Met Asn Tyr Thr Ala Tyr Tyr Pro Ser Pro Pro Gly Ser Pro
150         435             440             445
151 Asn Ser Leu Gly Tyr Phe Pro Thr Ala Ala Asn Leu Ser Gly Val Pro
152     450             455             460
153 Pro Gln Pro Gly Thr Val Val Arg Met Gln Gly Leu Ala Tyr Asn Thr
154 465             470             475             480
155 Gly Val Lys Glu Ile Leu Asn Phe Phe Gln Gly Tyr Gln Tyr Ala Thr
```

```
RAW SEQUENCE LISTING              DATE: 11/08/2000
PATENT APPLICATION:  US/09/697,206   TIME: 10:45:55

Input Set : A:\Seqlist.txt
Output Set: N:\CRF3\11082000\I697206.raw

156              485            490            495
157 Glu Asp Gly Leu Ile His Thr Asn Asp Gln Ala Arg Thr Leu Pro Lys
158              500            505            510
159 Glu Trp Val Cys Ile
160              515
162 <210> SEQ ID NO: 3
163 <211> LENGTH: 14
164 <212> TYPE: DNA
165 <213> ORGANISM: Synthetic
167 <400> SEQUENCE: 3
168 ttttgatcaa gctt                                        14
170 <210> SEQ ID NO: 4
171 <211> LENGTH: 42
172 <212> TYPE: DNA
173 <213> ORGANISM: Synthetic
175 <400> SEQUENCE: 4
176 ctaatacgac tcactatagg gctcgagcgg ccgcccgggc ag         42
178 <210> SEQ ID NO: 5
179 <211> LENGTH: 12
180 <212> TYPE: DNA
181 <213> ORGANISM: Synthetic
183 <400> SEQUENCE: 5
184 ggcccgtcct ag                                          12
186 <210> SEQ ID NO: 6
187 <211> LENGTH: 40
188 <212> TYPE: DNA
189 <213> ORGANISM: Synthetic
191 <400> SEQUENCE: 6
192 gtaatacgac tcactatagg gcagcgtggt cgcggccgag             40
194 <210> SEQ ID NO: 7
195 <211> LENGTH: 10
196 <212> TYPE: DNA
197 <213> ORGANISM: Synthetic
199 <400> SEQUENCE: 7
200 cggctcctag                                             10
202 <210> SEQ ID NO: 8
203 <211> LENGTH: 22
204 <212> TYPE: DNA
205 <213> ORGANISM: Synthetic
207 <400> SEQUENCE: 8
208 ctaatacgac tcactatagg gc                               22
210 <210> SEQ ID NO: 9
211 <211> LENGTH: 22
212 <212> TYPE: DNA
213 <213> ORGANISM: Synthetic
215 <400> SEQUENCE: 9
216 tcgagcggcc gcccgggcag ga                               22
218 <210> SEQ ID NO: 10
219 <211> LENGTH: 20
```

Per new Sequence Rules, the only valid responses are: Unknown, Artificial Sequence, or scientific name (Genus/species)

see circled portion of item 12 on Even summary sheet

```
220 <212> TYPE: DNA
221 <213> ORGANISM: Synthetic
223 <400> SEQUENCE: 10
224 agcgtggtcg cggccgagga                              20
226 <210> SEQ ID NO: 11
227 <211> LENGTH: 25
228 <212> TYPE: DNA
229 <213> ORGANISM: Homo Sapiens
231 <400> SEQUENCE: 11
232 atatcgccgc gctcgtcgtc gacaa                        25
234 <210> SEQ ID NO: 12
235 <211> LENGTH: 26
236 <212> TYPE: DNA
237 <213> ORGANISM: Homo Sapiens
239 <400> SEQUENCE: 12
240 agccacacgc agctcattgt agaagg                       26
242 <210> SEQ ID NO: 13
243 <211> LENGTH: 24
244 <212> TYPE: DNA
245 <213> ORGANISM: Synthetic
247 <400> SEQUENCE: 13
248 tcttgaaacc tccagacaca agaa                         24
250 <210> SEQ ID NO: 14
251 <211> LENGTH: 24
252 <212> TYPE: DNA
253 <213> ORGANISM: Synthetic
255 <400> SEQUENCE: 14
256 ggagatggta gacactggtg gagt                         24
258 <210> SEQ ID NO: 15
259 <211> LENGTH: 24
260 <212> TYPE: DNA
261 <213> ORGANISM: Synthetic
263 <400> SEQUENCE: 15
264 tcttgaaacc tccagacaca agaa                         24
266 <210> SEQ ID NO: 16
267 <211> LENGTH: 23
268 <212> TYPE: DNA
269 <213> ORGANISM: Synthetic
271 <400> SEQUENCE: 16
272 aagttacgat ttggcttcac tgg                          23
274 <210> SEQ ID NO: 17
275 <211> LENGTH: 9
276 <212> TYPE: PRT
277 <213> ORGANISM: Synthetic
279 <400> SEQUENCE: 17
280 Phe Leu Gly Glu Phe Ala Thr Asp Ile
281  1               5
283 <210> SEQ ID NO: 18
284 <211> LENGTH: 9
```

Please correct this error in subsequent sequences

VERIFICATION SUMMARY   DATE: 11/08/2000
PATENT APPLICATION: US/09/697,206   TIME: 10:45:56

Input Set : **A:\Seqlist.txt**
Output Set: **N:\CRF3\11082000\I697206.raw**

L:15 M:270 C: Current Application Number differs, Replaced Current Application No
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date